

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:58:25 ; Search time 16 Seconds  
(without alignments)  
623.987 Million cell updates/sec

Title: US-10-612-012-2

Perfect score: 4331

Sequence: 1 MSVSPQIPSRPLLLTHE.....AAPDGGFTVTRPGLNSKL 852

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.\*

1: /cgn2\_6/prodata/2/pubpa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/prodata/2/pubpa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/prodata/2/pubpa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/prodata/2/pubpa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1410	32.6	820	6	US-10-467-657-4910
2	1395.5	32.2	804	7	US-11-070-080-18
3	156	3.6	824	6	US-10-957-569-31
4	156	3.6	824	7	US-11-097-589-29
5	153.5	3.5	817	6	US-10-793-626-50
6	153.5	3.5	817	6	US-10-793-626-1528
7	150.5	3.5	611	6	US-10-454-437-54
8	150.5	3.5	697	6	US-10-485-517-202
9	150.5	3.5	925	6	US-10-454-437-50
10	144.5	3.3	716	6	US-10-131-826A-96
11	143	3.3	220	7	US-11-040-595-4
12	141.5	3.3	769	6	US-10-467-657-3280
13	141	3.3	655	6	US-10-467-657-3168
14	134.5	3.1	825	7	US-11-074-176-20
15	134	3.1	389	6	US-10-821-234-1401
16	133.5	3.1	1786	7	US-11-196-400-3
17	132	3.0	350	6	US-10-454-437-40
18	131	3.0	709	7	US-11-074-176-158
19	129.5	3.0	2665	7	US-11-124-368A-214
20	129.5	3.0	2668	7	US-11-124-368A-215
21	127.5	2.9	859	6	US-10-467-657-6084
22	126.5	2.9	488	6	US-10-485-517-307
23	125.5	2.9	459	6	US-10-467-657-3092
24	124.5	2.9	440	6	US-10-821-234-1330
25	124	2.9	1960	7	US-11-069-834-48

26 124 2.9 1960 7 US-11-069-834-50 Sequence 50, Appl  
27 123.5 2.9 426 6 US-10-454-437-70 Sequence 70, Appl  
28 123.5 2.9 426 6 US-10-454-437-72 Sequence 72, Appl  
29 122 2.8 420 7 US-11-074-176-48 Sequence 48, Appl  
30 120 2.8 406 6 US-10-821-234-1349 Sequence 1349, Ap  
31 120 2.8 406 6 US-10-878-558A-192 Sequence 192, App  
32 119 2.7 980 7 US-11-064-246-10 Sequence 10, Appl  
33 119 2.7 980 7 US-11-169-041-141 Sequence 141, App  
34 119 2.7 1652 6 US-10-995-561-663 Sequence 663, App  
35 119 2.7 1938 6 US-10-995-561-661 Sequence 661, App  
36 119 2.7 1938 6 US-10-995-561-662 Sequence 662, App  
37 119 2.7 1938 6 US-10-995-561-660 Sequence 660, App  
38 119 2.7 1972 6 US-10-995-561-664 Sequence 664, App  
39 119 2.7 1972 6 US-10-995-561-666 Sequence 666, App  
40 118 2.7 1268 6 US-10-995-561-918 Sequence 918, App  
41 118 2.7 1268 6 US-10-995-561-919 Sequence 919, App  
42 118 2.7 1268 6 US-10-995-561-920 Sequence 920, App  
43 116 2.7 989 6 US-10-821-234-975 Sequence 975, App  
44 115.5 2.7 2890 7 US-11-115-639-32 Sequence 32, Appl  
45 115 2.7 1618 6 US-10-984-645-2 Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-10-467-657-4910  
; Sequence 4910, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqMan99, version 1.04  
; SEQ ID NO 4910  
; LENGTH: 820  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-4910

Query Match 32.6%; Score 1410; DB 6; Length 820;  
Best Local Similarity 40.3%; Pred. No. 1.1e-94;  
Matches 311; Conservative 153; Mismatches 256; Indels 52; Gaps 12;  
QY 65 DPASDAQDPLRHRTGTAALAVQVGSNPKPHYTLITGLCRFOIVQVLKPKYPPIAEV 124  
DB 60 DAAVEEPVATDLYQTGTVAQVLQVL--KLPGTAKVLVEGLYRGRVLATIEDTGGFLVSHI 117  
QY 125 EOLDLEEFFPTCKRRELSELSEQFYKYAVQLVEMLDMSVPA--VAKLRLLDLSLPREA 182  
DB 118 EAVVEEDTGGNT-DLEAVRRTLLAQFQYA-----KLNKKIPAEITGSINGIAEN---SR 168  
QY 183 LPDILTSIRTSNKEKLQILDVAVLEERPKMTIPLLVRIQIEGLKLLQKTR---KPKQDD 239  
DB 169 LTDITVAHLQLKLAQRQILIPETGKMEFLAKLESELDIMOAEKIRGVRKQMEKS 228  
QY 240 KRVTAIRPIRITHISGTLDEDEDEDNDIVMLEKKIRTSMPQAHKVCVKIKRLKK 299  
DB 229 QREYVLN--EQIKAIHKELGEDE---NGELDALEAGIKAGMTKEASEKCLSELKCLKM 283  
QY 300 MPQSPPEVALTENVLELNVLPWNKSTTDRDIRAARILLNDNDHYAMEKLRVLYEYLAV 359  
DB 284 MPPMSAESTVVRNYIDTLLGLFWKKRSVSKAKAGLVLDADHYGLEKVRILEYLA 343

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:56:40 ; Search time 50 Seconds  
(without alignments)  
1408.794 Million cell updates/sec

Title: US-10-612-012-2

Perfect score: 4331

Sequence: 1 MSSVSPIQIPSRPLLLTHB.....AAPDGGFTVTRPGLLSKL 852

Scoring table: BLOSUM62

Gapop 10.0., Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pdp.\*
- 2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pdp.\*
- 3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pdp.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTRUS\_COMB.pdp.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pdp.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4331	100.0	852	2	US-09-741-150-2
2	4331	100.0	852	2	US-10-160-187-2
3	1869.5	43.2	884	2	US-09-741-150-4
4	1869.5	43.2	884	2	US-10-160-187-4
5	1745.5	40.3	810	2	US-09-902-340-11337
6	1407.5	32.5	820	2	US-09-252-991A-30204
7	1406.5	32.5	832	2	US-09-540-236-3056
8	1396.5	32.2	790	2	US-09-543-681A-6059
9	1390.5	32.1	822	2	US-09-328-352-5754
10	1378	31.8	845	2	US-09-489-039A-13590
11	1284	29.6	809	2	US-09-252-991A-29280
12	1251	28.9	937	1	US-08-253-155A-31
13	1251	28.9	937	2	US-09-538-092-1092
14	1251	28.9	937	2	US-09-949-002-374
15	1250	28.9	959	2	US-09-538-092-1091
16	1224	28.3	825	2	US-09-438-185A-29
17	1123	25.9	838	2	US-09-248-796A-17648
18	976	22.5	783	2	US-09-248-796A-16773
19	886	20.5	440	2	US-09-198-452A-44
20	407	9.4	562	2	US-09-949-002-426
21	390	9.0	458	2	US-09-270-767-46541
22	353	8.2	157	2	US-09-861-451A-8
23	336	7.8	424	2	US-09-198-452A-45
24	327	7.6	191	2	US-09-270-767-62133
25	165	3.8	3433	2	US-09-091-501B-10
26	165	3.8	3433	2	US-09-538-092-1136
27	162.5	3.8	900	2	US-09-248-796A-16848

28	161	3.7	761	2	US-09-328-352-7492	Sequence 7492, Ap
29	159	3.7	701	2	US-09-134-001C-3327	Sequence 3327, Ap
30	156.5	3.6	767	2	US-09-252-991A-31198	Sequence 31198, A
31	153.5	3.5	817	2	US-09-710-279-50	Sequence 50, Appl
32	153.5	3.5	817	2	US-09-710-279-1528	Sequence 1528, Ap
33	153.5	3.5	823	2	US-09-134-001C-4081	Sequence 4081, Ap
34	153.5	3.5	863	2	US-09-328-352-6730	Sequence 6730, Ap
35	152.5	3.5	726	2	US-09-252-991A-26767	Sequence 26767, A
36	152	3.5	422	2	US-09-134-001C-3034	Sequence 3034, Ap
37	151	3.5	335	2	US-09-134-000C-4044	Sequence 4044, Ap
38	150.5	3.5	611	2	US-09-602-777A-54	Sequence 54, Appl
39	150.5	3.5	697	2	US-08-816-177-2	Sequence 2, Appli
40	150.5	3.5	925	2	US-09-602-777A-50	Sequence 50, Appl
41	150	3.5	724	2	US-09-489-039A-12100	Sequence 12100, A
42	146.5	3.4	556	2	US-09-134-000C-3848	Sequence 3848, Ap
43	146	3.4	503	2	US-09-328-352-6479	Sequence 6479, Ap
44	145.5	3.4	689	2	US-09-134-000C-4499	Sequence 4499, Ap
45	145	3.3	546	2	US-09-489-039A-9627	Sequence 9627, Ap

ALIGNMENTS

RESULT 1  
US-09-741-150-2  
; Sequence 2, Application US/09741150  
; Patent No. 6436689  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CLO00968  
; CURRENT APPLICATION NUMBER: US/09/741,150  
; CURRENT FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 852  
; TYPE: PRT  
; ORGANISM: Human  
US-09-741-150-2

Query Match	100.0%	Score 4331	DB 2	Length 852
Best Local Similarity	100.0%	Pred. No. 0		
Matches	852	Conservative	0	Mismatches
			0	Indels
			0	Gaps
QY	1	MSSVSPIQIPSRPLLLTHB	VS	SAHNLQVRSRLKGTSLQSTILGVI
DB	1	MSSVSPIQIPSRPLLLTHB	VS	SAHNLQVRSRLKGTSLQSTILGVI
QY	61	PNTDPASDAQDLPPLHRI	GT	AAALVQVGVSNPKPHYTLITGLCRFQIVQVLEKPY
DB	61	PNTDPASDAQDLPPLHRI	GT	AAALVQVGVSNPKPHYTLITGLCRFQIVQVLEKPY
QY	121	IAEVEQLDRLSEPPNCKR	RE	LGELSEQFYKAVOLVEMLDMSVPAVKRLLSLPR
DB	121	IAEVEQLDRLSEPPNCKR	RE	LGELSEQFYKAVOLVEMLDMSVPAVKRLLSLPR
QY	181	EALPDILTITISNKEKQI	LD	AVSLERFKMTIPLLVRQIEGLKLLQKTRPKQDDDK
DB	181	EALPDILTITISNKEKQI	LD	AVSLERFKMTIPLLVRQIEGLKLLQKTRPKQDDDK
QY	241	RVIATIRPIRITHISGTL	DE	DEDEDDNDIVMLEKKIRTSMPQAHKVCVKIKKLM
DB	241	RVIATIRPIRITHISGTL	DE	DEDEDDNDIVMLEKKIRTSMPQAHKVCVKIKKLM
QY	301	QSMPEYALTRYLELMVEL	P	NKSTTDRLDRAARILLNDNDHYAMEKLRVLEAVR
DB	301	QSMPEYALTRYLELMVEL	P	NKSTTDRLDRAARILLNDNDHYAMEKLRVLEAVR
QY	361	QLKNLKGPILCFVGPGV	G	KTSVGRSVAKTILGRBPHRIALGVCQSDIRGHRRTYGS
DB	361	QLKNLKGPILCFVGPGV	G	KTSVGRSVAKTILGRBPHRIALGVCQSDIRGHRRTYGS

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:48:40 ; Search time 259 Seconds  
(without alignments)  
2320.888 Million cell updates/sec

Title: US-10-612-012-2

Perfect score: 4331

Sequence: 1 MSSVSPIQPSRLPLLTHE.....AAFDGGFTVKTRPGLLSKL 852

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	4323	99.8	852	2	Q86W48	HUMAN	Q86W48 homo sapien
2	4319	99.7	852	2	Q8CE9	HOMO SAPIEN	Q8CE9 homo sapien
3	4292	99.1	852	2	QSR6M5	PONPY	Q5x6m5 pongo pygma
4	4154	95.9	852	2	Q3DBN5	MOUSE	Q9dbn5 m mus muscu
5	3629	83.8	715	2	Q8N3B9	HUMAN	Q8n3b9 homo sapien
6	3376	77.9	840	2	QSPQY6	BRARE	Q5pqy6 brachydario
7	2972	68.6	581	2	Q96K43	HUMAN	Q96k43 homo sapien
8	2863	66.1	806	2	Q6TX13	RAT	Q6tx13 rattus norv
9	2089	48.2	432	2	Q8BK80	MOUSE	Q8bk80 mus musculu
10	1890.5	43.7	874	2	Q4T681	TETNG	Q4t681 tetraodon n
11	1869.5	43.2	885	1	LN0H1	MAIZE	P93647 zea mays (m
12	1861	43.0	884	2	Q94F60	9POAL	Q94f60 dichantheli
13	1850	42.7	886	2	QSR597	WHEAT	Q6rs97 triticum ae
14	1849	42.7	888	1	LN0H1	ARATH	O64948 arabidopsis
15	1846.5	42.6	884	2	Q6GV57	ORYSA	Q6gv57 oryza sativ
16	1812.5	41.8	880	2	Q69SH2	ORYSA	Q69sh2 oryza sativ
17	1801	41.6	875	1	LN0H1	SPIOL	O04979 spinacia ol
18	1784	41.2	843	2	LN0P66	9DELTA	Q4npw6 anaeromykob
19	1745.5	40.3	826	1	LN02	MYXXA	P36774 myxococcus
20	1596	36.9	819	2	Q74752	GEOSL	Q74752 geobacter s
21	1585.5	36.6	788	2	Q82V32	NITEU	Q82v32 nitrosonoma
22	1573	36.3	775	2	Q5KWK1	GEOKA	Q5kwk1 geobacillus
23	1561	36.0	774	2	Q5GJ36	BACLD	Q65gj6 bacillus li
24	1559.5	35.0	776	2	Q817Q4	BACCR	Q817q4 bacillus ce
25	1546.5	35.7	796	2	Q9H21	BRABJA	Q89h21 bradyrhizob
26	1545.5	35.7	776	2	Q6HSS4	BACAN	Q6hss4 bacillus an
27	1545	35.7	557	2	Q8GT60	ARATH	Q8gt60 arabidopsis
28	1544.5	35.7	776	2	Q633X4	BACC2	Q633x4 bacillus ce
29	1543.5	35.6	773	2	Q81LC1	BACAN	Q81lc1 bacillus ce
30	1541.5	35.6	776	2	Q6HDS6	BACHK	Q6hd56 bacillus th
31	1540	35.6	856	2	Q72CU2	DESVH	Q72cu2 desulfovibr

32 1535.5 35.5 773 2 Q4MS75 BACCCE Q4ms75 bacillus ce  
33 1535.5 35.5 773 2 Q72ZV6 BACC1 Q72zv6 bacillus ce  
34 1535 35.4 772 2 Q8CXB9 OCEIH Q8cxb9 oceanobacill  
35 1525.5 35.2 774 2 Q9K3F6 BACRD Q9k3f6 bacillus na  
36 1525 35.2 783 2 Q7MMZ2 BORBR Q7mmz2 bordetella  
37 1525 35.2 783 2 Q7WBH0 BORPA Q7wbh0 bordetella  
38 1509 34.8 778 2 Q8RC23 THETN Q8rc23 thermoanaer  
39 1506 34.8 779 1 LN0 BRECH LN0 brech  
40 1498 34.6 775 2 Q5WE91 BACSK Q5we91 brevibacill  
41 1497 34.6 774 1 LN01 BACSU LN01 bacillus su  
42 1493 34.5 779 2 Q84FG5 9BACL Q84fg5 brevibacill  
43 1482 34.2 768 2 Q74EN9 GEOSL Q74en9 geobacter s  
44 1477.5 34.1 932 2 Q4WVD9 ASPFU Q4wvd9 aspergillus  
45 1476.5 34.1 937 2 Q7SA85 NEUCR Q7sa85 neurospora

## ALIGNMENTS

RESULT 1  
Q86W48\_HUMAN  
ID Q86W48\_HUMAN PRELIMINARY; PRT; 852 AA.  
AC Q86W48;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Peroxisomal lon protease.  
GN Name=LONP;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA De Walque S., Van Veldhoven P.P.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madañ A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield A.S.N., Krzywinska M.I., Skaleka U., Smailus D.B.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RG NIH MGC Project;  
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
RA EMBL; AJ548761; CAD68987.1; -; mRNA.  
DR EMBL; BC093912; AAH93912.1; -; mRNA.  
DR EMBL; BC093910; AAH93910.1; -; mRNA.  
DR HSP; P08177; IRR9.  
DR MEROPS; S16.006; -.  
DR Ensembl; ENSG00000102910; Homo sapiens.  
DR GO; GO:0005524; P:ATP binding; IEA.

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:52:15 ; Search time 48 Seconds

(without alignments)  
1707.848 Million cell updates/sec

Title: US-10-612-012-2

Perfect score: 4331

Sequence: 1 MSVSPIQIPSRPLLLTHE.....AADFGGFTVKTRPGLNSKL 852

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1869.5	43.2	885	1 T04321	endopeptidase La h
2	1801	41.6	875	2 T09142	endopeptidase La h
3	1745.5	40.3	827	1 A36895	endopeptidase La (
4	1525.5	35.2	774	2 B84031	ATP-dependent prot
5	1506	34.8	779	1 B42375	endopeptidase La (
6	1497	34.6	774	1 I40421	endopeptidase La (
7	1473.5	34.0	817	1 A49844	endopeptidase La (
8	1467	33.9	821	2 B75530	ATP-dependent prot
9	1453.5	33.6	813	2 G75331	ATP-dependent prot
10	1445.5	33.4	787	2 A72230	endopeptidase La (
11	1433	33.1	820	2 H81908	probable endopepti
12	1426	32.9	820	2 H81106	ATP-dependent prot
13	1406.5	32.5	798	2 G83420	Lon protease PAL
14	1403.5	32.4	786	2 G82141	ATP-dependent LA p
15	1400.5	32.3	799	2 C87492	endopeptidase La (
16	1398	32.3	795	1 A70322	endopeptidase La (
17	1393	32.2	778	2 E97224	ATP-dependent Lon
18	1372.5	31.7	784	2 AE0558	Lon protease (impo
19	1372	31.7	823	2 AF3361	endopeptidase La (
20	1371.5	31.7	779	2 E97778	endopeptidase La (
21	1369.5	31.6	805	2 AH2731	ATP-dependent prot
22	1369.5	31.6	805	2 G97512	ATP-dependent ser1
23	1368	31.6	848	2 C82712	ATP-dependent ser1
24	1367.5	31.6	784	1 SUECLA	endopeptidase La (
25	1367	31.6	810	1 JEC045	endopeptidase La (
26	1365.5	31.5	784	1 A71704	endopeptidase La (
27	1364	31.5	803	1 A64070	endopeptidase La (
28	1362.5	31.5	784	2 E90690	endopeptidase La (
29	1358	31.4	799	2 A85541	hypothetical prote

30	1355	31.3	784	2 AC0383	endopeptidase La (
31	1351.5	31.2	786	2 A96956	ATP-dependent prot
32	1350	31.2	784	1 S47270	endopeptidase La (
33	1349	31.1	777	2 P84985	endopeptidase La (
34	1309.5	30.2	773	2 T27382	hypothetical prote
35	1301.5	30.1	964	1 T04325	probable ATP-depen
36	1293.5	29.9	831	1 A71825	endopeptidase La (
37	1289.5	29.8	881	1 B71316	endopeptidase La (
38	1286	29.7	799	2 P83549	probable ATP-depen
39	1282.5	29.6	835	1 C64692	endopeptidase La (
40	1259.5	28.1	791	2 D81310	endopeptidase La (
41	1251	28.9	937	1 S42366	endopeptidase La h
42	1244	28.7	819	2 E81681	protease, Lon fa
43	1242	28.7	819	1 C71527	endopeptidase La (
44	1238	28.6	962	1 S57342	endopeptidase La h
45	1235.5	28.5	813	1 D70176	endopeptidase La (

#### ALIGNMENTS

##### RESULT 1

T04321

endopeptidase La homolog (EC 3.4.21.-) Lon1 precursor, mitochondrial - maize  
N:Alternate names: ATP-dependent proteinase LON1; ATP-dependent serine proteinase La h  
C:Species: Zea mays (maize)  
C:Dates: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04321

R:Barakat, S.; Pearce, D.A.; Sherman, F.; Rapp, W.D.

Plant Mol. Biol. 37, 141-154, 1998

A:Title: Maize contains a Lon protease gene that can partially complement a yeast pim1-

A:Reference number: Z15282; MUID:98281582; PMID:9620272

A:Accession: T04321

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-885 <BAR>

A:Cross-references: UNIPROT:P93647; UNIPARC:UPI0000043401; EMBL:U85494; NID:ig1816585; P:

C:Genetics:

C:Gene: LON2

C:Function:

C:Description: serine proteinase

C:Superfamily: ATP-dependent Lon protease

C:Keywords: ATP; DNA binding; hydrolase; mitochondrial matrix; mitochondrion; molecular

P:409-416/Region: nucleotide-binding motif A (P-loop)

F:472-477/Region: nucleotide-binding motif B

F:781/Active site: Ser #status predicted

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

Query Match 43.2%; Score 1869.5; DB 1; Length 885;

Best Local Similarity 44.7%; Pred. No. 3.1e-98;

Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:48:20 ; Search time 200 Seconds

(without alignments)  
1871.754 Million cell updates/sec

Title: US-10-612-012-2

Perfect score: 4331

Sequence: 1 MSSVSFPIQPSRLPLLTLE.....AADFQGFVTKTRPGLNSKL 852

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_21.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*  
9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4331	100.0	852	5	ABG32363 Human ATP
2	4331	100.0	852	7	ADM56266 Human ATP
3	4331	100.0	852	9	ADW85873 Human pro
4	4323	99.8	852	7	AAb47562 Protease
5	4323	99.8	852	7	ADC10040 Human NOV
6	4319	99.7	852	4	AAm93785 Human pol
7	4319	99.7	852	8	ADL31771 Human pro
8	4316	99.7	852	5	ABP69443 Human pol
9	3520	81.3	693	7	ADJ70423 Human hea
10	2972	68.6	581	4	AAb95332 Human pro
11	2972	68.6	581	4	AAb99179 Human ATP
12	2957	68.3	581	8	ADR08963 Human pro
13	2921	67.4	571	4	AAb94288 Human pro
14	2201	50.8	432	4	AAb97167 Human ATP
15	2201	50.8	432	4	AAE11085 Human ATP
16	2165	50.0	433	6	ABU11611 Human MDD
17	1983.5	45.8	423	4	AAU17375 Novel sig
18	1983.5	45.8	423	7	ADB94083 Human nov
19	1974	45.6	391	5	ABb89248 Human pol
20	1869.5	43.2	884	9	ADM56268 Human ATP
21	1869.5	43.2	884	9	ADW85875 Zea maye
22	1849	42.7	887	8	ADN74069 Thale cre
23	1840.5	42.5	888	8	Adt60761 Plant pol
24	1745.5	40.3	810	9	ABm92138 M. xanthu

25	1745.5	40.3	827	8	ADS30399	Ads30399 Bacterial
26	1585.5	36.6	770	8	ADS30692	Ads30692 Bacterial
27	1557.5	36.0	769	8	ADS27589	Ads27589 Bacterial
28	1546.5	35.7	794	6	ABU17842	Abu17842 Protein e
29	1540.5	35.6	767	8	ADN27141	Adn27141 Bacterial
30	1540	35.6	786	8	ADN26262	Adn26262 Bacterial
31	1525.5	35.2	774	8	ADS28353	Ads28353 Bacterial
32	1497	34.6	774	8	ADS44794	Ads44794 Bacterial
33	1470.5	34.0	817	8	ADS30184	Ads30184 Bacterial
34	1468.5	33.9	771	8	ADS21365	Ads21365 Bacterial
35	1467	33.9	821	8	ADS44986	Ads44986 Bacterial
36	1461	33.7	777	8	ADN26487	Adn26487 Bacterial
37	1453.5	33.6	813	8	ADN17992	Adn17992 Bacterial
38	1450	33.5	788	8	ADS22145	Ads22145 Bacterial
39	1446.5	33.4	773	6	ABU24441	Abu24441 Protein e
40	1445.5	33.4	787	8	ADN20397	Adn20397 Bacterial
41	1433	33.1	820	6	ABU37980	Abu37980 Protein e
42	1426	32.9	789	6	ABU25481	Abu25481 Protein e
43	1426	32.9	820	8	ADP08282	Adp08282 Neisseria
44	1415.5	32.7	804	6	ABU39533	Abu39533 Protein e
45	1414.5	32.7	816	9	ABE411304	Abe411304 L. pneumo

#### ALIGNMENTS

##### RESULT 1

ABG32363  
ID ABG32363 standard; protein; 852 AA.

XX AC ABG32363;

XX DT 15-NOV-2002 (first entry)

XX DE Human ATP-dependent protease.

XX KW Human; enzyme; ATP-dependent protease; developmental disorder; cancer;

XX KW retinoblastoma; melanotic melanoma; endometrial adenocarcinoma;

XX KW ovarian adenocarcinoma; schizophrenia.

XX OS Homo sapiens.

XX PN US2002081704-A1.

XX PD 27-JUN-2002.

XX PF 21-DEC-2000; 2000US-00741150.

XX PR 22-NOV-2000; 2000US-0252410P.

XX PA (GUEG/) GUEGLER K.

XX PA (WEBB/) WEBSTER M.

XX PA (YANC/) YAN C.

XX PA (SHAO/) SHAO W.

XX PA (KSTC/) KETCHUM K A.

XX PA (DFRA/) DI FRANCESCO V.

XX PA (BEAS/) BEASLEY B M.

XX PI Guegler K, Webster M, Yan C, Shao W, Ketchum KA, Di Francesco V;

XX PI Beasley EM;

XX DR WPI; 2002-635469/68

XX DR N-FSDB; ABK90887; ABK90888.

XX PT New human ATP-dependent peptides and encoding nucleic acids, useful for

XX PT diagnosing, preventing and/or treating disorders like cancer, e.g.

XX PT retinoblastoma, melanoma, endometrial and ovarian adenocarcinomas, and

XX PT schizophrenia.

XX PS Claim 1; Fig 2; 119pp; English.

XX CC The invention relates to a new isolated ATP-dependent protease (I) and

XX CC the nucleic acid encoding it. The peptides are useful in identifying

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:57:30 ; Search time 177 Seconds  
(without alignments)  
2011.246 Million cell updates/sec

Title: US-10-612-012-2

Perfect score: 4331

Sequence: 1 MSSVSPQIPSLRPLLLTHE.....AAPDGGFTVKTTRGLNSKL 852

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA Main:
- 1: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pap.\*
  - 2: /cgn2\_6/prodata/1/pubpa/US08\_PUBCOMB.pap.\*
  - 3: /cgn2\_6/prodata/1/pubpa/US09\_PUBCOMB.pap.\*
  - 4: /cgn2\_6/prodata/1/pubpa/US10\_PUBCOMB.pap.\*
  - 5: /cgn2\_6/prodata/1/pubpa/US10B\_PUBCOMB.pap.\*
  - 6: /cgn2\_6/prodata/1/pubpa/US11\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4331	100.0	852	US-09-741-150-2	Sequence 2, Appl1
2	4331	100.0	852	US-10-612-012-2	Sequence 2, Appl1
3	4323	99.8	852	US-10-239-032-4	Sequence 4, Appl1
4	3520	81.3	693	US-10-408-765A-2229	Sequence 2229, Ap
5	2201	50.8	432	US-10-156-239-44	Sequence 44, Appl
6	1983.5	45.8	423	US-09-764-868-940	Sequence 940, App
7	1974	45.6	391	US-10-264-237-1624	Sequence 1624, Ap
8	1869.5	43.2	884	US-09-741-150-4	Sequence 4, Appl1
9	1869.5	43.2	884	US-10-612-012-4	Sequence 4, Appl1
10	1869.5	43.2	885	US-10-156-239-47	Sequence 47, Appl
11	1840.5	42.5	887	US-10-739-930-10838	Sequence 10838, A
12	1745.5	40.3	827	US-10-369-493-19432	Sequence 19432, A
13	1657.5	38.3	850	US-10-437-963-156016	Sequence 156016,
14	1585.5	36.6	770	US-10-369-493-19725	Sequence 19725, A
15	1557.5	36.0	769	US-10-369-493-16222	Sequence 16222, A
16	1546.5	35.7	794	US-10-282-122A-45766	Sequence 45766, A
17	1540.5	35.6	767	US-10-369-493-9794	Sequence 9794, Ap
18	1540.5	35.6	786	US-10-369-493-17386	Sequence 8915, Ap
19	1525.5	35.2	774	US-10-369-493-17386	Sequence 17386, A
20	1522	35.1	624	US-10-767-701-44549	Sequence 44549, A
21	1497	34.6	774	US-10-369-493-23224	Sequence 23224, A
22	1470.5	34.0	817	US-10-369-493-19217	Sequence 19217, A
23	1468.5	33.9	771	US-10-369-493-10398	Sequence 10398, A
24	1467	33.9	821	US-10-369-493-23416	Sequence 23416, A
25	1461	33.7	777	US-10-369-493-9140	Sequence 9140, Ap
26	1453.5	33.6	813	US-10-369-493-645	Sequence 645, App
27	1450	33.5	788	US-10-369-493-11178	Sequence 11178, A

#### ALIGNMENTS

##### RESULT 1

US-09-741-150-2

Sequence 2, Application US/09741150

Publication No. US20020081704A1

GENERAL INFORMATION:

APPLICANT: GUEGLER, Karl et al

TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: CLO00968

CURRENT APPLICATION NUMBER: US/09/741,150

CURRENT FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 852

TYPE: PRT

ORGANISM: Human

US-09-741-150-2

Query Match	100.0%	Score 4331	DB 3	Length 852
Best Local Similarity	100.0%	Pred. No. 0		
Matches	852	Conservative	0	Mismatches
			0	Indels
			0	Gaps
			0	
QY	1	MSSVSPQIPSLRPLLLTHEGVLLPGSTWRTSVDSAHNLQVRSRLKGTSLQSTILGVI	60	
DB	1	MSSVSPQIPSLRPLLLTHEGVLLPGSTWRTSVDSAHNLQVRSRLKGTSLQSTILGVI	60	
QY	61	PNTDPPASDAQDLPPLHRIPTAALAVQVVGSNWPKPHYTLITGLCRFOIVQLKPKYP	120	
DB	61	PNTDPPASDAQDLPPLHRIPTAALAVQVVGSNWPKPHYTLITGLCRFOIVQLKPKYP	120	
QY	121	IAEVQLDRLEFPNTCKMRELGSLSEQFYKAVQVLEMDMSVPAVAKRLRLDLSLR	180	
DB	121	IAEVQLDRLEFPNTCKMRELGSLSEQFYKAVQVLEMDMSVPAVAKRLRLDLSLR	180	
QY	181	BALPDILTSIIRTSNKEKLIQILDVSLERFNTIPLLVQRIEGLKLLQKTRPKQDDDK	240	
DB	181	BALPDILTSIIRTSNKEKLIQILDVSLERFNTIPLLVQRIEGLKLLQKTRPKQDDDK	240	
QY	241	RVIAIRPIRRIITHISGTTLEDEDEDDNDIVMLEKKIRTSSMPEQAHKVCVKEIKLKKM	300	
DB	241	RVIAIRPIRRIITHISGTTLEDEDEDDNDIVMLEKKIRTSSMPEQAHKVCVKEIKLKKM	300	
QY	301	POSMEPEYALTRYLSELMVELPWKSTTDRLDRAARILLNDNDHYANBKLRVLEYLAVR	360	
DB	301	POSMEPEYALTRYLSELMVELPWKSTTDRLDRAARILLNDNDHYANBKLRVLEYLAVR	360	
QY	361	QLKNNKGPILCFVGGPGVGKTSVGRSVAKTIGREFPHRIALGVCVQDSQDIRGHRRTYVGS	420	